Jurnal Inovasi Ekonomi
Vol. 08 No. 01 April 2023 Pages: 1-10
P-ISSN: 2477-4804 E-ISSN: 2686-3804 http://ejournal.umm.ac.id/index.php/jiko

GRDP forecasting through night-time light data: Evidence from Indonesia

Agung Andiojaya[1*]; R.R. Nefriana[2]; Riana Rizka[3]; Titi Kanti Lestari[4]

Statistician, Central Bureau of Statistic-Indonesia, Indonesia[1,2,3,4]

Received: 17-02-2021 | Revision: 23-03-2022 | Accepted: 28-03-2022 | Published: 26-12-2022
DOI: https://doi.org/10.22219/jiko.v8i01.20239

Abstract
Gross Regional Domestic Product (GRDP) is a widely used macroeconomic indicator to measure socioeconomic development. A comprehensive GRDP calculation requires complete data available from the national level down to minor administrative areas. However, institutions that carry out GRDP calculations often face challenges in obtaining complex data at smaller administrative levels. This paper explores the use of nighttime light (NTL) data as an alternative data source to estimate GRDP and its changes. The results of the panel data regression with random effects show that NTL data is statistically significant in describing GRDP in various Indonesian provinces. However, during the pandemic, NTL was not substantial in describing GRDP and showed movements that were not in line with GRDP.

Keywords: Night-time light; GRDP; Indonesia

*Corresponding author: andiojaya@stis.ac.id

1. Introduction

Gross Regional Domestic Product (GRDP) is a widely used macroeconomic indicator to measure socioeconomic development (Bilan et al., 2020; Elshendy & Fronzetti Colladon, 2017; Grishin et al., 2019; Marikina, 2018). It represents the sum of the value-added produced by all regional production units or the total value of final goods and services (net) made by all economic units (BPS, 2014). To obtain a comprehensive GRDP calculation, Statistics Indonesia requires complete data from the national level down to minor administrative areas.

However, Statistics Indonesia faces difficulties in obtaining data for smaller administrative levels, as the effort required to obtain such data is complex and challenging. For example, low response rates from household and business respondents persist over time (BPS, 2017, 2018, 2019, 2021). Additionally, local governments often struggle to provide detailed data, making it burdensome to allocate national accounts. Furthermore, taxation data obtained by other National Statistical Offices from taxation offices is not yet implemented in Indonesia (Luisa et al., 2020; Statistics, 2004).

Due to a taxation law that protects individual data, detailed data such as annual revenue or output cannot be drawn from governmental agencies like Statistics Indonesia. To fill the data gap, National Accounts conduct ad-hoc surveys, but resource constraints and the non-availability of Subject Matters Areas responsible for covering some economic industries often lead to insufficient sample sizes and purposive sampling (Suharni & Pertiwi, 2016). The non-availability of these Subject Matter Areas also requires Statistics Indonesia to estimate some required numbers for financial, real estate, business, and other services economic industries. These data gap challenges are becoming more significant during the pandemic, as traditional data collection approaches, especially surveys, are not viable and arduous (Suharni & Pertiwi, 2016).

At the same time as the challenges mentioned earlier, an opportunity emerges. Some success stories show the potential of using satellite imagery data as a proxy for some economic indicators. In 2012, Henderson, Storeygard, and Weil reported that they successfully utilized observed night-time light (NTL) as an indicator of income across countries and sub- and supranational regions (Henderson et al., 2012). Prakash, Shukla, Bhowmick, and Beyer (2019) also found that night-time light correlates strongly with GDP and other critical macroeconomic indicators such as industrial production and credit growth at a national level in India. Furthermore, in 2020, Adhikari and Dhital revealed that decentralization hindered regional convergence between first and second subnational regions within a country by exploiting night-time lights captured by the U.S. Air Force satellites combined with fiscal, political, and administrative decentralization databases (Adhikari & Dhital, 2020).

Considering this empirical success, we investigated the cross-section correlation between the NTL intensities measured by satellites from outer space and levels of GRDP in Indonesia. By approximating GRDP by night light data, we hope to alleviate the data collection challenges mentioned earlier. Our study contributes to an alternative approach to GRDP projections, especially in developing countries like

Indonesia, while simultaneously challenging previous findings that state that NTL is a robust proxy.

2. Method

This study analyzes data from 33 provinces in Indonesia quarterly, starting from Quarter-1 2018 to Quarter-4 2020. The data used in this study include Indonesia's GRDP at a constant price by province sourced from BPS-Statistics Indonesia. Then, for NTL satellite data, we employ the average day-night-band (DNB) values of Monthly Cloud-free DNB Composite NPP-VIIRS data sourced from Earth Observation Group. To obtain spatial data for each province in Indonesia, this study uses Indonesian geometric data in the form of Global Administrative Unit Layers (GAUL) and First-Level Administrative Units sourced from FAO UN.

We utilize Google Earth Engine (GEE) to generate NTL data per province. GEE is a cloud-based platform for planetary-scale geospatial analysis, which allows the processing of various geographical data at scale and handles large geographical datasets. GEE provides access to numerous remotely sensed datasets and derived products, including VIIRS DNB and Global Administrative Unit Layers (Gorelick et al., 2017).

This research aims to evaluate the effectiveness of NTL as the alternative variable to predict GRDP. Considering that each province's GRDP data has varying values and trends, it is not enough to analyze the trend using univariate time series (Andiojaya, 2021). To shed light on this matter, this study uses panel data regression to get a comprehensive outlook on the variable interaction (Croissant & Millo, 2008; Gujarati, 2012; Hill et al., 2018). Generally, the panel data analysis model can be expressed in equation (1).

$$Log(GDRP_{it}) = \alpha + \beta\, Log(NTL_{it}) + \varepsilon_{it} \tag{1}$$

Note: $GDRP_{it}$ is GRDP at Constant Price of i^{th} province and t^{th} quarter; α is intercept; β is regression coefficient; NTL_{it} is nighttime-light Data of i^{th} province and t^{th} quarter; ε_{it} = error term of i^{th} province and t^{th} quarter; i = 1 to N; t = 1 to T

Before developing the appropriate empirical model for NTL and GRDP, the panel data model must be ascertained first. Therefore, the first step in this research is to determine which model is more suitable for the data. According to Ekananda (2018) and Nachrowi & Usman (2006), the selection of the panel data model can be based on the purpose of the analysis or due to various mathematical technical problems. Since this research focuses on the effect of the independent variable on the dependent variable by considering individual characteristics (province's value) and building the assumption of the uncorrelated unobservable variable to the independent variable, the random effect method is theoretically sound for this data. Several tests were conducted to provide robust evidence for this assumption and to determine the appropriate panel data effect method. The Hausman and Breusch-Pagan Lagrange Multiplier tests were carried out (Baltagi & Li, 1991; Ekananda, 2018; Greene, 2012; Nachrowi & Usman, 2006).

In addition to investigating the relationship between NTL and GRDP in general, this paper also explores the relationship between NTL data and GRDP before and during the pandemic. This study generates a binary dummy variable to distinguish the quarter before and after the pandemic. This dummy variable marks the period from 2018 Q1 to 2019 Q4 as the period before the COVID-19 pandemic and serves as the baseline for the dummy variable. Thus, the model built in this paper follows equation (2).

$$Log(GDRP_{it}) = \alpha + \beta\, Log(NTL_{it}) + DummyPandemic_t + DummyPandemic_t * Log(NTL_{it}) + \varepsilon_{it} \quad (2)$$

3. Empirical Result

The results of the Haussman and Breusch-Pagan Lagrange Multiplier tests reported in Table 1 indicate that the random effect model is better than the fixed effect and common effect models. The Haussman test result suggests a random effect since the test did not exhibit enough statistical evidence to reject the null hypothesis. Similarly, the Breusch-Pagan Lagrange Multiplier test shows enough statistical evidence to reject the null hypothesis, leading to the selection of the random effect model as the preferred model.

Table 1. Hypothesis and significance result of panel data effect test

Test	Hypothesis	*p*-value
Haussman	H_0: using Random Effect H_a: using Fixed Effect	0.975
Breusch-Pagan Lagrange Multiplier	H_0: using Common Effect H_a: using Random Effect	0.000

Evaluation of NTL as Predictor for GRDP

Table 2 presents three key findings. First, the model's optimistic estimator of Log NTL and its statistical significance indicate that NTL can be statistically used to estimate changes in GRDP. A positive estimator value suggests that an increase in light intensity will lead to an increase in GRDP, and vice versa. Second, while the p-value of the DummyPandemic variable is not statistically significant, the negative value of the estimator may indicate that the GRDP value during the pandemic tended to be lower than before the pandemic. This finding confirms that the Covid-19 pandemic caused a decline in economic growth in Indonesia throughout 2020. Third, the estimator value of the interaction results of the LogNTL and DummyPandemic variables confirms the second finding. The estimator of the LogNTL and DummyPandemic variables is -0.012, indicating that the relationship between nightlight growth and GDP growth during the pandemic was 0.012 points lower than before the pandemic.

One important finding that needs to be addressed is the low R^2. This coefficient of determination indicates that the model can only explain around 11 percent of the variation in the dependent variable. According to (Cohen 1992) and Field (2018), this R^2 value suggests a low interaction between the regressor and dependent variable. Despite the model's small R^2, the overall model's significance indicates that NTL is still a promising predictor variable for GRDP. Interestingly, when the model generated data during the pandemic (Q1-2020 to Q4-2020), it was found to be insignificant, and the adjusted R^2 was small. This model suggests that during the pandemic, NTL cannot be used as a proxy for measuring GRDP.

Table 2. Random effect model

Variable(s)	Estimate	Std.Error	Prob.
Intercept	17.5338	0.1548	0.000 ***
Log_NTL	0.0201	0.0032	0.000***
DummyPandemic	-0.0030	0.0054	0.583
Log_NTL * DummyPandemic	-0.0124	0.0036	0.00057 ***
R^2			0.1083
Adjusted R^2			0.1015
Chi Square on 3 DoF			47.631
p-value			0.000

Note on Signif. 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 3. Random effect model during pandemic

Variable(s)	Estimate	Std.Error	Prob.
Intercept	17.5181	0.1656	0.000 ***
Log_NTL	-0.0085	0.0057	0.1365
R^2			0.0168
Adjusted R^2			0.0092
Chi Square on 3 DoF			2.2172
p-value			0.1365

Note on Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Figure 1 presents the real value of GRDP, while Figure 2 shows the predicted value of GRDP calculated based on the model. The two figures demonstrate that the GRDP values in some areas exhibit similar movements with the observed GRDP values, such as in West Sumatra, Riau, Jambi, Bengkulu, and North Sulawesi. However, for other provinces, the predicted value of GRDP generated by the model does not show promising results.

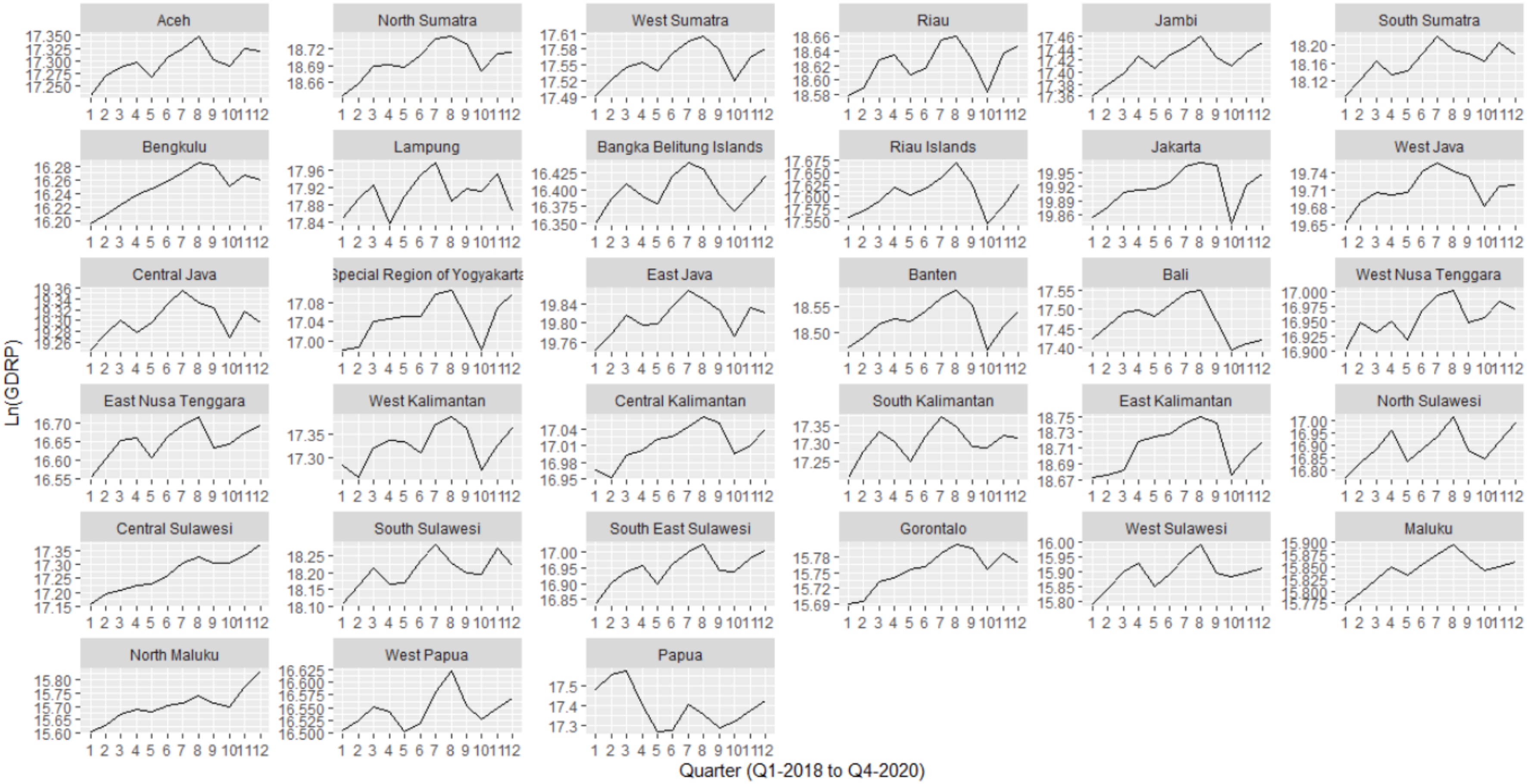

Figure 1. Trend line of observed GRDP by Province from Q1-2018 to Q4-2020

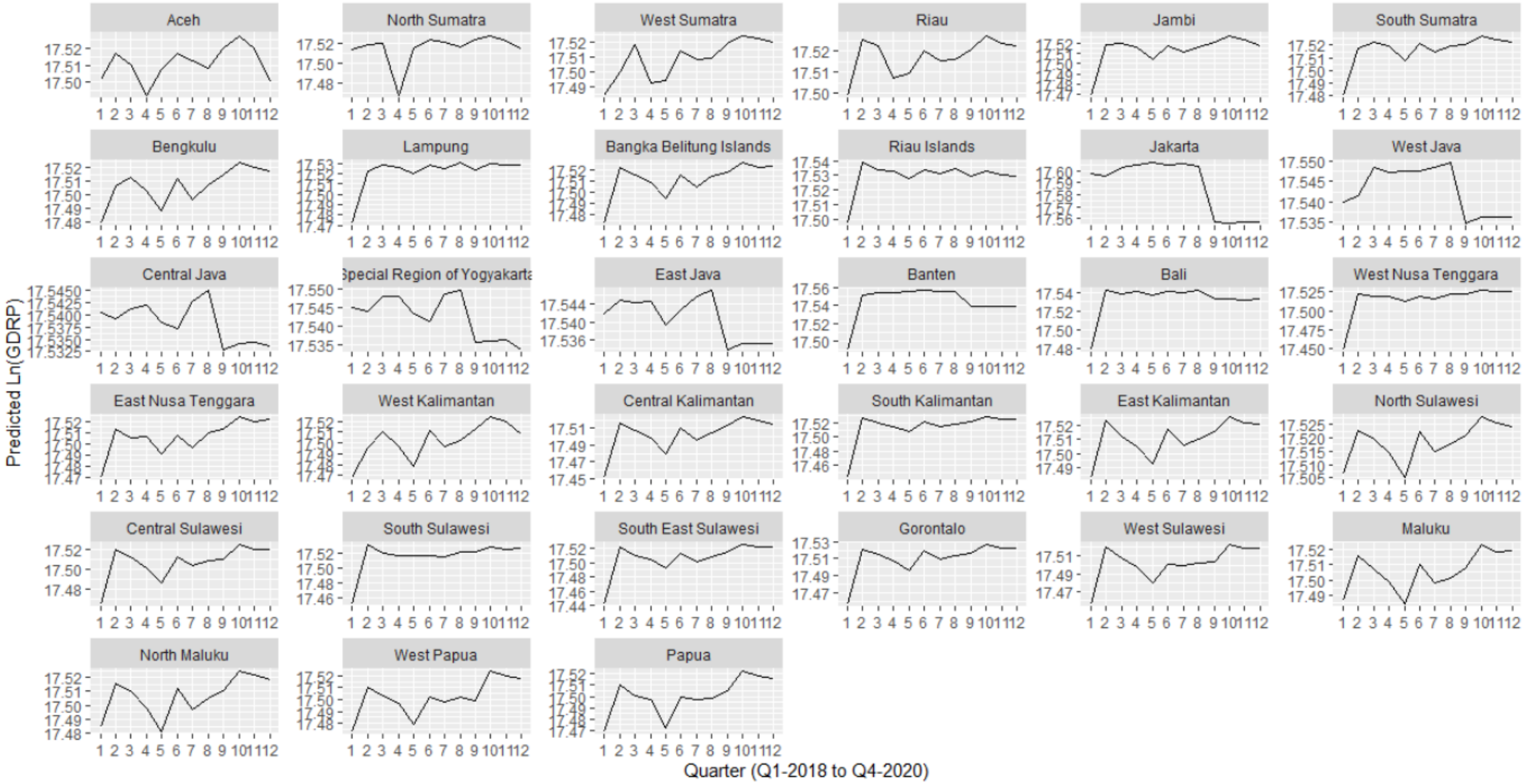

Figure 2. Trend line of predicted GRDP by Province from Q1-2018 to Q4-20

4. Conclusions

Based on the model developed in this study, the NTL variable is found to be statistically significant in explaining variations in the value of GRDP. The model provides a good predicted value of GRDP whose trend is similar to the observed value of GRDP in several provinces, but in other provinces, the predicted value tends to move in a different direction from the actual value. During the pandemic, the model shows an insignificant relationship between NTL and GRDP. From the author's perspective, this result might be due to the dominance of the informal sector over the formal sector in Indonesia's economic structure. Therefore, future research needs to consider other data as control variables to overcome the limitations of this model. Despite its shortcomings, NTL can still be considered as an alternative data source for measuring a region's economy.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study

5. References

Adhikari, B., & Dhital, S. (2020). *Decentralization and Regional Convergence: Evidence from Night-Time Lights Data. Economic Inquiry, 59*(3). https://doi.org/10.1111/ecin.12967.

Andiojaya, A. (2021). Transmisi Harga Gabah terhadap Harga Beras: Tinjauan Arah, Besaran dan Lama Perubahan. *JSEP (Journal of Social and Agricultural Economics), 14*(2). https://doi.org/10.19184/jsep.v14i2.24304

Baltagi, B. H., & Li, Q. (1991). A joint test for serial correlation and random individual effects. *Statistics & Probability Letters*, *11*(3). https://doi.org/10.1016/0167-7152(91)90156-L

Bilan, Y., Mishchuk, H., Roshchyk, I., & Kmecova, I. (2020). An analysis of intellectual potential and its impact on the social and economic development of European countries. *Journal of Competitiveness*, *12*(1). https://doi.org/10.7441/joc.2020.01.02

Badan Pusat Statistik. (2014). *Sistem Neraca Nasional 2008* (1st ed.). Badan Pusat Statistik. https://bps.go.id/publication/2008/02/18/d686364900b131e819f08724/sistem-neraca-nasional-2008.html

Badan Pusat Statistik (2017). *Laporan Kinerja BPS Tahun 2016*. Badan Pusat Statistik. https://www.bps.go.id/publication/2017/12/25/a9872cc44091bf8d4106eaca/laporan-kinerja-badan-pusat-statistik-tahun-2016.html

Badan Pusat Statistik. (2018). *Laporan Kinerja BPS Tahun 2017*. Badan Pusat Statistik. https://ppid.bps.go.id/upload/doc/LAKIN_Deputi_Bidang_Statistik_Sosial_Tahun_2018_1658474816.pdf

Badan Pusat Statistik. (2019). *Laporan Kinerja BPS Tahun 2018*. Badan Pusat Statistik. https://ppid.bps.go.id/upload/doc/LAKIN_Inspektorat_Utama_2018_1658475364.pdf

Badan Pusat Statistik. (2021). *Laporan Kinerja BPS Tahun 2020*. Badan Pusat Statistik.

https://ppid.bps.go.id/upload/doc/Laporan_Kinerja_BPS_2020_1658385891.pdf

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1). https://doi.org/10.1037/0033-2909.112.1.155

Croissant, Y., & Millo, G. (2008). Panel Data Econometrics in R : The plm Package. *Journal of Statistical Software*. https:// 10.18637/jss.v027.i02

Ekananda, M. (2018). Analisis Ekonometrika Data Panel Edisi 2: Teori Lengkap Dan Pembahasan Menyeluruh Bagi Peneliti Ekonomi, Bisnis, Dan Sosial (2nd ed.). *Mitra Wacana Media*. https://scholar.ui.ac.id/en/publications/analisis-ekonometrika-data-panel-edisi-2-teori-lengkap-dan-pembah

Elshendy, M., & Fronzetti Colladon, A. (2017). Big data analysis of economic news: Hints to forecast macroeconomic indicators. *International Journal of Engineering Business Management*, *9*, 1–12. https://doi.org/10.1177/1847979017720040

Field, A. P. (2018). Discovering statistics using IBM SPSS statistics. *Sage Publications Inc*. https://doi/10.5555/2502692

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*. https://doi.org/10.1016/j.rse.2017.06.031

Greene, W. H. (2012). Econometric analysis (7th ed). *Prentice Hall*. https://abs/10.1080/10543406.2012.701589

Grishin, V. I., Ustyuzhanina, E. V., & Komarova, I. P. (2019). Main Problems with Calculating GDP as An Indicator of Economic Health of The Country. *International Journal of Civil Engineering and Technology (IJCIET)*, *10*(2).

Gujarati, D. (2012). Econometrics by Example. https://link.springer.com/book/9781137375018

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, *102*(2). https://doi.org/10.1257/aer.102.2.994

Hill, R. C., Griffiths, W. E., & Lim, G. C. (2018). Principles of Econometrics (5th Ed). Wiley.

Luisa, R., Thompson, C., & Jones, J. (2020). A Statistical Business Register Spine as a New Approach to Support Data Integration and Firm-level Data Linking: An ABS Perspective. *Statistical Journal of the IAOS*, *1*. https://doi/10.3233/SJI-200640

Marikina, M. (2018). Gross Domestic Product or Gross National Happiness—Which is The Better Alternative for Economic Measurement?. *34th International Scientific Conference on Economic and Social Development - XVIII International Social Congress (ISC-2018)*, 188. https://ideas.repec.org/a/nwe/natrud/y2019i5p175-186.html

Nachrowi, N. D., & Usman, H. (2006). Pendekatan Populer dan Praktis Ekonometrika untuk Analisis Ekonomi dan Keuangan (1st ed.). *Lembaga Penerbit Fakultas Ekonomi Universitas Indonesia*. https://lib.ui.ac.id/detail.jsp?id=108874

Prakash, A., Shukla, A. K., Bhowmick, C., Carl, R., & Beyer, M. (2019). Night-time Luminosity: Does it Brighten Understanding of Economic Activity in India?. *Reserve Bank of India Occasional Papers*, *40*(1). https://www.researchgate.net/publication/334811462_Night-time_Luminosity_Does_it_Brighten_Understanding_of_Economic_Activity_in_India

Suharni, L., & Pertiwi, R. P. (2016). Harmonization and Integration of Internal Data Sources in Constructing Statistical Business Register (SBR) Initial Data in BPS - Statistics Indonesia. *Asia-Pacific Economic Statistics Week Seminar Component*. https://doi/10.23889/ijpds.v6i1.1680